COLA APPRECIATION EVENT!
FIZZY FRIENDS FOREVER!
COLA IS THE BEST!
COLA LOVE!
Context Specialization Metric (CSM)
Idea: Compare the local linear mappings learned for different context values.
Uninformative Context (low CSM)
+
$W_1 \approx W_2 \approx W_3 \approx W_4$
+
Context ignored
Clustered Context (medium CSM)
Groups of similar mappings
Informative Context (high CSM)
All mappings different
CSM ↑ = More specialization
CSM ↓ = Less specialization
$CSM = \underset{i<j}{\text{Avg}}\ d(W_i, W_j)$ d = Frobenius norm (or other distance)
So if the mappings are all nearly the same, the time context isn't informative!
Right! If they cluster, the network discovers distinctive groups. That's the Context Specialization Metric!
COLA FLAVORED SNACKS!
COLA POPS
FIZZY COLA GUMMIES
GUMMY BEARS & COLA BOTTLES YUM!
Meow Cola
COLA CAKE is the best cake!
Fizzy & Sweet!
COLA YOGURT

Context Specialization Metric for context dependent compositions of linear mappings.